

Kvantitativne metode u psihologiji



KANONIČKA KORELACIONA ANALIZA

Šta je kanonička korelaciona analiza?



- Metod multivarijacione analize koji se koristi za utvrđivanje postojanja veze i jačine povezanosti dva skupa promjenljivih.
- Korelacija je uobičajeno sredstvo analize povezanosti dvije promjenljive
- Stepenn međusobne povezanosti jedne zavisne i više nezavisnih varijabli – koeficijent višestruke korelacije
- Kad imamo više zavisnih promjenljivih – kanonička korelaciona analiza

O idejama kanoničke korelacione analize



- Originalni teorijski doprinos razvoju dao je Hotelling (1936. godine)
- KKA – postupak za ispitivanje povezanosti dva skupa varijabli
- Šta nam to omogućava?
- Zašto bi nas zanimala povezanost između dva skupa varijabli?
 - Nije uvijek moguće izmjeriti direktno pojavu koja nas interesuje (mjerimo preko skupa dobrih indikatora), npr. porodični odnosi
 - Teško zaključivanje na bazi ogromnog broja pojedinačnih koef.korelacije
 - Statistički neispravno testirati stat.značajnost velikog broja koeficijenata korelacije jer ti stat.testovi ne bi bili nezavisni
 - Za analizu povezanosti dva skupa moraju se uzeti u obzir interkorelacije

O kanoničkoj korelacionoj analizi



- Standardno sredstvo u statističkoj analizi
- Mnogobrojni primjeri kanoničke korelacije u ekonomiji, psihologiji, medicini, meteorologiji itd.

Primjer:

- Polazimo od pretpostavke da su odnosi u porodici povezani sa konceptom o sebi (self-concept) i želimo to da ispitamo.
- Porodične odnose možemo mjeriti preko nekih varijabli (odnos sa majkom, ocem, toplina, neke karakteristike interakcije...)
- Koncept o sebi (kroz npr. set individualnih atributa i karakteristika ličnosti o sebi)

Kanonička korelacija



Prosta korelacija $\longrightarrow Y_1 = X_1$

Višestruka korelacija $\longrightarrow Y_1 = X_1 \quad X_2 \quad X_3$

Kanonička korelacija $\longrightarrow Y_1 \quad Y_2 \quad Y_3 = X_1 \quad X_2 \quad X_3$

Pojednostavljeni postupak iza KKA



- Počínje se sa skupom Y varibli i X varijabli

$$y_1 \ y_2 \ y_3 = x_1 \ x_2 \ x_3$$

- Konstruišu se kanoničke varijable kao linearna kombinacija Y varijabli

$$CV_{y1} = b_1 y_1 + b_2 y_2 + b_3 y_3$$

- Konstruišu se kanoničke varijable kao linearna kombinacija X varijabli

$$CV_{x1} = b_1 x_1 + b_2 x_2 + b_3 x_3$$

- Kanočka korelacija je korelacija kanoničkih varijabli

$$R_c = r_{cvy1, cvx1}$$

Pojednostavljeni postupak iza KKA



- Kako ona to radi?

- Nalaženjem linearnih kombinacija varijabli, iz jednog odnosno drugog skupa, koje su u što je moguće većoj korelaciji (znači tako da korelacije tih varijabli dobijenih linearnim kombinovanjem budu maksimalne).
- Varijable koje se dobiju takvim kombinovanjem nazivaju se kanoničke varijable (kanoničke komponente ili kanonički faktori).
- Kanoničke varijable se proizvode u parovima (jedna iz jednog i jedna iz drugog skupa), a parova može da ima najviše onoliko koliko ima varijabli u manjem od dva skupa.

Zadaci kanoničke korelacije



- Utvrdi jačinu veze koja postoji između dva skupa varijabli
- Formira linearne kombinacije unutar skupa zavisnih i posebno unutar skupa nezavisnih varijabli, ali tako da između te dvije linearne kombinacije postoji maksimalna korelacija
- Na osnovu tih linearnih kombinacija možemo proniknuti u prirodu veze koja postoji među varijablama
- Umjesto ispitivanja većeg broja originalnih promjenljivih, izučavamo vezu između izvedenih promjenljivih
- Izdvajanje parova linearnih kombinacija

Ograničenja i uslovi za primjenu kanoničke analize



- Za korišćenje u deskriptivne svrhe nema pretpostavki o rasporedu slučajnih vektora, a mogu se koristiti i promjenljive na nominalnoj ili ordinalnoj skali
- Međutim, statistički postupak zaključivanja o značajnosti veze, kao i značajnost kanoničke korelacije zahtijeva uvođenje pretpostavke o višedimenzionalnom rasporedu.
- Tehnika je osjetljiva na outlier-e
- Da bi rezultati bili stabilni, potrebno je da broj ispitanika bude 40-60 puta veći od broj varijabli

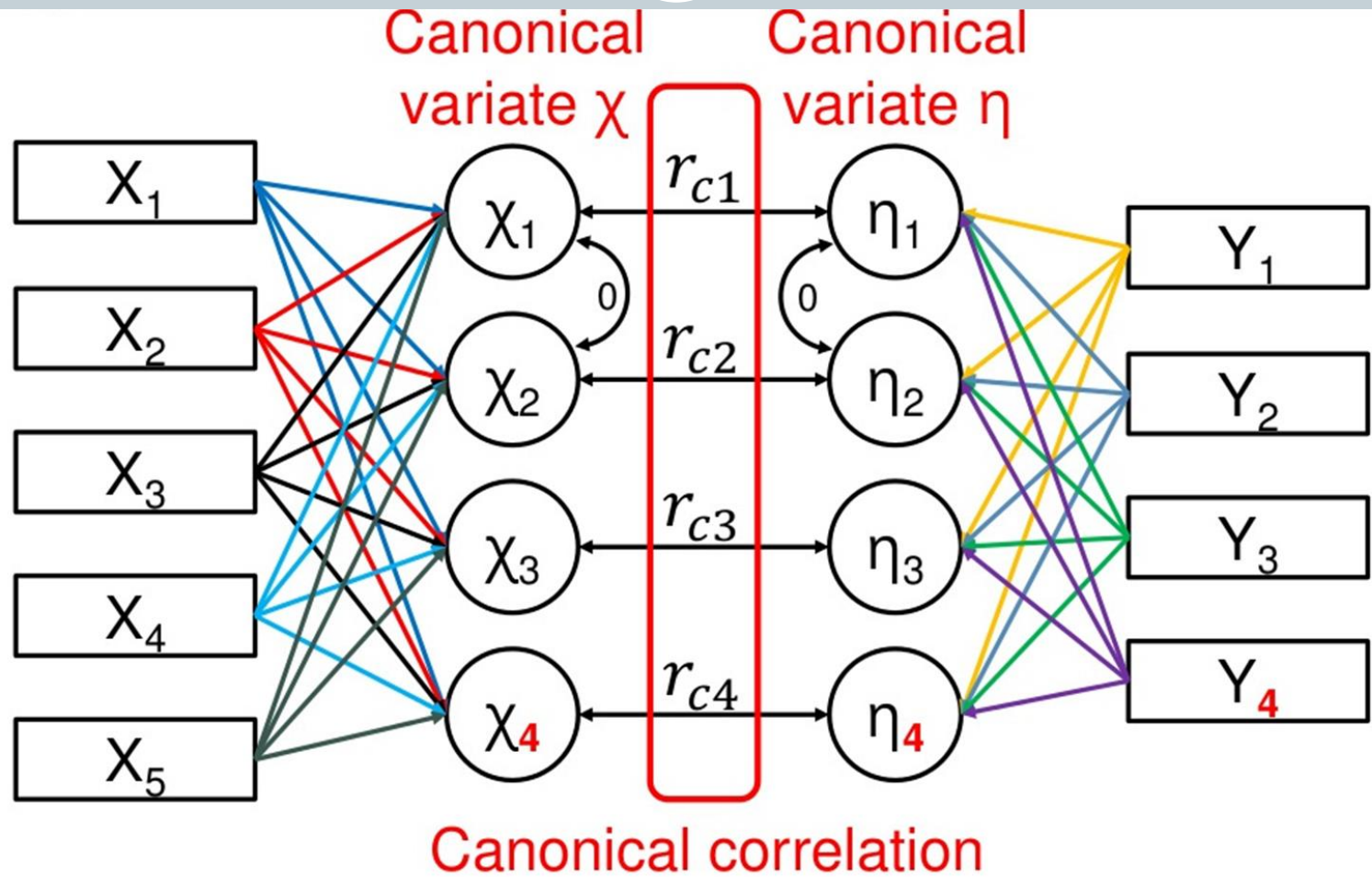
Dodatna pojašnjenja



- **Faktorska vs kanonička korelaciona analiza**
 - Kod FA se faktori prave tako da objašnjavaju što je moguće veću količinu varijanse, dok je ovdje fokus na tome da kanoničke varijable (faktori) budu u što je moguće većoj međusobnoj korelaciji
 - Moguće je da se desi da ono što je glavni izvor međusobne povezanosti varijabli iz jednog skupa nije istovremeno i izvor njihove povezanosti sa drugim skupom.

Neka ograničenja KKA: interpretabilnost, linearna veza, senzitivnost, nestandardne opservacije itd.

Kanonička analiza



Glavni rezultati



- **Koeficijent kanoničke korelacije** – koeficijent korelacije dvije kanoničke komponente. Ona u stvari ne predstavlja realnu već prije neku maksimalnu moguću povezanost dva skupa varijabli.
- **Koeficijent kanoničke determinacije** – kvadrat koeficijenta kanoničke korelacije – varijansa koja je zajednička dvijema kanoničkim komponentama
- **Koeficijent strukture** (engl. canonical loadings) – koeficijent korelacije određene izvorne varijable sa kanoničkom komponentom dobijenom iz skupa varijabli kome ta izvorna varijabla pripada.
- **Koeficijent krosstrukture** (engl. cross loadings) – koeficijent korelacije određene izvorne varijable sa kanoničkom komponentom dobijenom iz onog drugog skupa varijabli (kome ta varijabla ne pripada).

Glavni rezultati



- **Sopstvena vrijednost** (karakteristična vrijednost, karakteristični koren, eigenvalue) – količina varijanse koju varijable jednog skupa dijele sa kanoničkom komponentom dobijenom iz tog skupa
- **Objašnjena varijansa** – proporcija ukupne varijanse izvornih varijabli koju one dijele sa datom kanoničkom varijablom (dobijena iz istog skupa kao izvorne varij.)
- **Koeficijent redundanse** – količina varijanse koju varijable iz jednog skupa dijele sa određenom kanoničkom komponentom iz drugog skupa. Redundansa nam detaljnije govori o povezanosti između dva skupa varijabli.
- **Kanonički koeficijenti** (nalaze se u matrici koeficijenata) – konstante sa kojima treba množiti izvorne varijable da bi se dobila određena kanonička komponenta. Oni praktično pokazuju kako se došlo do kanoničkih varijabli (tj. kako su one formirane iz izvornih).

O proceduri...



- Testira se statistička značajnost svake od dobijenih kanoničkih korelacija, tako da ima smisla interpretirati samo one kanoničke parove koji su u stat.značajnoj korelaciji
- Prvi kanonički par je u najvećoj međusobnoj korelaciji, drugi u manjoj i treći u još manjoj itd.
- Onda je zgodno uzeti u obzir one kanoničke parove koji visoko međusobno koreliraju, a zanemariti one koji koreliraju malo.
- Ipak stvar zavisi prije svega od materije koja se ispituje i od za tu materiju relevantnih teorijskih stvari (šta očekujemo i želimo da dobijemo).

Primjer u SPSS-u



Radimo na skupu od 600 podataka sa osam posmatranih varijabli. 3 varijable su psihološke (mjesto kontrole, samo-koncept i motivacija), a 4 varijabli opisuje akademska postignuća (preko standardiziranih testova čitanja, pisanja, matematike i nauke). Uz to, varijabla “ženski pol” je varijabla koja uzima vrijednost 0 i 1. Jedinica je indikator da se radi o studentkinji.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
locus of control	600	-2.23	1.36	.0965	.67028
self-concept	600	-2.62	1.19	.0049	.70551
motivation	600	.00	1.00	.6608	.34273
reading score	600	28.30	76.00	51.9018	10.10298
writing score	600	25.50	67.10	52.3848	9.72645
math score	600	31.80	75.50	51.8490	9.41474
science score	600	26.00	74.20	51.7633	9.70618
female	600	.00	1.00	.5450	.49839
Valid N (listwise)	600				

Statistics

female

N	Valid	600
	Missing	0

female

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	273	45.5	45.5	45.5
	1.00	327	54.5	54.5	100.0
Total		600	100.0	100.0	

[illegible][illegible]

Primjer u SPSS-u



Eigenvalues and Canonical Correlations

Root No.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	.27450	87.33628	87.33628	.46409	.21538
2	.02887	9.18537	96.52164	.16751	.02806
3	.01093	3.47836	100.00000	.10399	.01081

Dimension Reduction Analysis

Roots	Wilks L.	F	Hypoth. DF	Error DF	Sig. of F
1 TO 3	.75436	11.71573	15.00	1634.65	.000
2 TO 3	.96143	2.94446	8.00	1186.00	.003
3 TO 3	.98919	2.16461	3.00	594.00	.091

Primjer u SPSS-u



EFFECT .. WITHIN CELLS Regression (Cont.)

Univariate F-tests with (5,594) D. F.

Variable	Sq. Mul. R	Adj. R-sq.	Hypoth. MS	Error MS	F	Sig. of F
locus_of	.18062	.17372	9.72160	.37123	26.18789	.000
self_con	.01957	.01131	1.16669	.49212	2.37076	.038
motivati	.07874	.07098	1.10799	.10913	10.15338	.000

Raw canonical coefficients for DEPENDENT variables

Function No.

Variable	1	2	3
locus_of	1.25383	-.62148	.66169
self_con	-.35135	-1.18769	-.82672
motivati	1.26242	2.02726	-2.00023

Standardized canonical coefficients for DEPENDENT variables

Function No.

Variable	1	2	3
locus_of	.84042	-.41656	.44352
self_con	-.24788	-.83793	-.58326
motivati	.43267	.69480	-.68554

Primjer u SPSS-u



Correlations between DEPENDENT and canonical variables
Function No.

Variable	1	2	3
locus_of	.90405	-.38969	.17562
self_con	.02084	-.70874	-.70516
motivati	.56715	.35089	-.74513

Variance in dependent variables explained by canonical variables

CAN. VAR.	Pct Var DEP	Cum Pct DEP	Pct Var COV	Cum Pct COV
1	37.97982	37.97982	8.17994	8.17994
2	25.90966	63.88948	.72701	8.90694
3	36.11052	100.00000	.39050	9.29745

Raw canonical coefficients for COVARIATES
Function No.

COVARIATE	1	2	3
read	.04462	-.00491	-.02138
write	.03588	.04207	-.09131
math	.02342	.00423	-.00940
science	.00503	-.08516	.10984
female	.63212	1.08464	1.79465

Primjer u SPSS-u



Standardized canonical coefficients for COVARIATES
CAN. VAR.

COVARIATE	1	2	3
read	.45080	-.04961	-.21601
write	.34896	.40921	-.88810
math	.22047	.03982	-.08848
science	.04878	-.82660	1.06608
female	.31504	.54057	.89443

Correlations between COVARIATES and canonical variables
CAN. VAR.

Covariate	1	2	3
read	.84045	-.35883	-.13536
write	.87654	.06484	-.25456
math	.76395	-.29795	-.14776
science	.65841	-.67680	.23036
female	.36411	.75493	.54340

Variance in covariates explained by canonical variables

CAN. VAR.	Pct Var DEP	Cum Pct DEP	Pct Var COV	Cum Pct COV
1	11.30458	11.30458	52.48769	52.48769
2	.70132	12.00590	24.99409	77.48177
3	.09804	12.10394	9.06617	86.54795

* SPSS nema “point-and-click” (pokaži kursorom i klikni) opciju za kanoničku korelacionu analizu već se ona izvodi korišćenjem nekog makroa praljenog za tu namenu.